DOCUMENT RESUME

ED 137 395 TM 006 201

AUTHOR Molina, Huberto; Shoemaker, David M.
TITLE Comprehensive Assessment of Spanish Language Proficiency Using Multiple Matrix Sampling Techniques.
PUB DATE [Apr 77]
NOTE 9p.; Paper presented at the Annual Meeting of the American Educational Research Association (61st, New York, New York, April 4-8, 1977); For a related document, see ED 093 911

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS *Bilingual Students; English (Second Language); *Item Sampling; *Language Proficiency; *Measurement Techniques; Norm Referenced Tests; Spanish; *Spanish Speaking; Statistical Analysis; *Student Testing

ABSTRACT

In discussion of assessment and its applicability to the needs of the Spanish-speaking student population, four types of assessments are presented in the context of the educationally significant testing needed in this area. The focus of this paper is on the use of comprehensive assessment to measure the repertoire of language resources that the Spanish-speaking population bring to the school experience. Multiple matrix sampling techniques are described as they apply to the following assessment requirement: To what degree do speakers of Spanish in the United States possess proficiency in Spanish, taking into account levels of formality, slow/fast speech, colloquial, and dialectal usage of language? (Author)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED137395

Comprehensive Assessment of Spanish Language Proficiency Using Multiple Matrix Sampling Techniques

HUBERTO MOLINA
SWRL Educational Research and Development

DAVID M. SHOEMAKER
U.S. Office of Education

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Paper presented at the American Educational Research Association Annual Meeting, April 4-8, 1977, New York City

TM006 201

Comprehensive Assessment of Spanish Language Proficiency Using Multiple Matrix Sampling Techniques

HUBERTO MOLINA
SWRL Educational Research and Development

DAVID M. SHOEMAKER
U.S. Office of Education

In discussing assessment and its applicability to the needs of the Spanish speaking student population, it behoves one to examine testing in the context of needed educationally significant instruments in this area. Four types of language assessment are discussed in this paper with emphasis on Multiple Matrix Sampling.

Built-in assessment instruments providing individual scores

Preinstructional assessment is used to identify children who will profit from a particular instructional program. The assessment score is also used to suggest placement of a child in an instructional sequence. Built-in assessments provide enroute information to identify instructional areas where individual remediation is needed. End of program and end of year assessment are used to assess the degree to which individual children have mastered program outcomes. Built-in assessment scores provide information on an individual basis and on a class, district, state, or national level can be valuable in tracking student performance within the program and beyond the program. For a detailed presentation of analyses using built-in assessment scores, see Molina et al., 1975.

Micro testing

Micro testing involves a small number of subjects used to test a restricted number of variables, e.g., the acquisition of negation by bilingual children. Results from these tests are useful for selecting and sequencing items in constructing tests and instructional programs.

Norm-reference testing

Norm-referenced assessment provides information concerning the total population across the nation. The in-depth information needed for norm-referenced assessment leads to the next type of assessment.

Comprehensive assessment

An in-depth assessment should address questions of the following type. To what degree do speakers of Spanish in the United States possess listening, speaking, reading, and writing skills taking into account levels of formality, slow/fast speech, colloquial, and dialectal use of language. As part of such a comprehensive assessment, various socio-linguistic contexts would have to be used to assure that natural language situations include home, neighborhood, school, and work domains. To illustrate, within the home domain interactions of the following type should be included: child-parent and child with younger/older siblings. Topics should include usual every day conversation and include references to holidays and other special events which are an integral part of the life experience of Spanish speaking children. This type of assessment information provided by a comprehensive Spanish language assessment coupled with an English language assessment of similar scope can provide directions to communities by which to make curriculum decisions. Instructional programs could then be selected taking into account language and cultural resources that children bring to the school experience. Assessment carried out periodically can be used to describe the changing linguistic competencies of the learner population.

Method of testing. Comprehensive testing across domain and skill areas necessarily results in a large pool of items and a large number of students to be tested. Assessment procedures call for division of the item population into item subsets or subtests with each subtest administered to a

subgroup of students selected from those being tested. The large number of test items and students necessitates that different students respond to different subtests. The results, however, provide an estimate of how all students would have performed over all the items in the item population. To accomplish the assessment task, multiple matrix sampling techniques apply to the assessment of language proficiency over a wide array of language items.

The scope of language assessment is such that the items to be included will exceed the number that can be administered to a given child. Of chief concern is how all students would perform on all the item population. Although it is not possible to administer all items to all students, it is possible to estimate through multiple matrix sampling the results which would have been obtained had this been done.

Multiple Matrix Sampling. In multiple matrix sampling the item population is divided through random sampling into item subsets or subtests with each subtest administered to a subgroup of students selected randomly from the student population. In this procedure, different students respond to different subtests but the results may be used to estimate how all students would have performed over all the items in the item population. The advantages of multiple matrix sampling over alternative procedures are discussed in detail by Shoemaker (1973). The chief advantages of the procedures are the following:

1. A more comprehensive and detailed assessment is possible because it is student performance on an item population--instead of, for example, a standardized test or some item subset--which is

being determined. Were any item aggregate used in place of this item population, it is quite possible that the results would be specific to that particular test.

2. Because different students respond to different subtests (and not all students to one large item aggregate), a sharp decrease in testing time over traditional testing procedures is afforded.
3. In multiple matrix sampling, it is not necessary for all students to be tested. Both items and students are sampled in this procedure.
4. Multiple matrix sampling is also suited ideally to periodic testing of student performance because subtests are assigned randomly to students on both occasions. This avoids the problems (e.g., the carry-over) which may occur when the same test is administered on both occasions.
5. Estimates of student performance obtained from multiple matrix sampling are more accurate on the average than those obtained from any item subset.

Multiple Matrix Sampling has been used in the context of Spanish/English bilingual education. An investigation was conducted to develop a procedure by which student performance over the linguistic components in both languages could be assessed in a rigorous and comprehensive manner. The purpose of that investigation was to determine the strengths and weaknesses of the instructional program so that specific recommendations could be made concerning how best to expand the bilingual program. The investigation was successful in establishing a procedure for testing linguistic components in an efficient and inexpensive manner and the results obtained indicate

clearly the relative emphasis in each content area needed to bring about bilingual performance. For a description of the evaluation of the Bilingual Spanish/English Program see Molina and Shoemaker 1973.

Multiple Matrix Sampling as a measurement of group performance has been adopted for the assessment of student performance at the state level in California and Florida. A detailed explanation of how California used the procedure to assess mathematics achievement is given by Tardiff (1971) and Riles (1971).

Multiple Matrix Sampling can be used in another bilingual/bicultural assessment context. Used in conjunction with a relevant item universe, it provides an excellent procedure for examining the relative merits of competing instructional programs. This approach to the problem stems from recognition that the primary measurement requirement in evaluating competing programs is group assessment and not the assessment of individual differences among students. The key feature in this approach is developing an item universe which spans broadly the content domain under consideration. Such an item universe will always include more content and more objectivity than that considered within any individual program. To accomplish this, the evaluation procedure should attempt to assess the broadest spectrum of objectives and domains. In comparing programs in which objectives are not clearly specified, items must be inferred.

In discussing assessment and its applicability to the needs of the Spanish-speaking student population several different types of testing approaches are needed. The focus of this paper has been on the use of comprehensive assessment offered by Multiple Matrix Sampling techniques to measure the repertoire of language resources that the Spanish-speaking

population brings to the school experience and its use in examining the relative merits of competing programs.

References

Molina, Huberto, et al. Development of the SWRL English Language and Concepts Program for Spanish-Speaking Children (LCS), Monograph, SWRL Educational Research and Development, Los Alamitos, California, 1975.

Shoemaker, David. Principles and Procedures of Multiple Matrix Sampling. Ballinger Publishing Company. Cambridge, Massachusetts, 1973.

Molina, H. and Shoemaker, D. A preliminary evaluation of a bilingual Spanish/English program using multiple matrix sampling. ERIC Document Reproduction Service, P. O. Drawer O, Bethesda, Maryland 20014, Order Number: ED093-911.

Riles, W. Final Report of Test Development of the Miller Mathematics Improvement Project. Sacramento, California: California State Department of Education, 1971.

Tardiff, R. F. Report of the November 1971 SCIMA Test Results. California State Department of Education, Office of Program Evaluation, 1971.